THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2012-02

A randomized gossip consensus algorithm on convex metric spaces

Ion Matei, Christoforos Somarakis, John S. Baras

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2012**	2. REPORT TYPE	3. DATES COVERED **00-00-2012 to 00-00-2012**
4. TITLE AND SUBTITLE **A randomized gossip consensus algorithm on convex metric spaces**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland, College Park,Institute for Systems Research,2173 A.V. Williams Building,College Park,MD,20742**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S)
		11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
A consensus problem consists of a group of dynamic agents who seek to agree upon certain quantities of interest. This problem can be generalized in the context of convex metric spaces that extend the standard notion of convexity. In this paper we introduce and analyze a randomized gossip algorithm for solving the generalized consensus problem on convex metric spaces. We study the convergence properties of the algorithm using stochastic differential equations theory. We show that the dynamics of the distances between the states of the agents can be upper bounded by the dynamics of a stochastic differential equation driven by Poisson counters. In addition, we introduce instances of the generalized consensus algorithm for several examples of convex metric spaces together with numerical simulations.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**20**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

A randomized gossip consensus algorithm on convex metric spaces

Ion Matei, Christoforos Somarakis, John S. Baras

Abstract

A consensus problem consists of a group of dynamic agents who seek to agree upon certain quantities of interest. This problem can be generalized in the context of convex metric spaces that extend the standard notion of convexity. In this paper we introduce and analyze a randomized gossip algorithm for solving the generalized consensus problem on convex metric spaces. We study the convergence properties of the algorithm using stochastic differential equations theory. We show that the dynamics of the distances between the states of the agents can be upper bounded by the dynamics of a stochastic differential equation driven by Poisson counters. In addition, we introduce instances of the generalized consensus algorithm for several examples of convex metric spaces together with numerical simulations.

I. Introduction

Distributed algorithms are found in applications related to sensor, peer-to-peer and ad-hoc networks. A particular distributed algorithm is the *consensus* (or agreement) algorithm, where a group of dynamic agents seek to agree upon certain quantities of interest by exchanging information among them, according to a set of rules. This problem can model many phenomena involving information exchange between agents such as cooperative control of vehicles, formation control, flocking, synchronization, parallel computing, etc. Distributed computation over networks has a long history in control theory starting with the work of Borkar and Varaiya [1], Tsitsikils, Bertsekas and Athans [25], [26] on asynchronous agreement problems and parallel computing. A theoretical framework for solving consensus problems was introduced by Olfati-Saber and Murray in [16], [17], while Jadbabaie et al. [7] studied alignment problems for reaching an agreement. Relevant extensions of the consensus problem were done by Ren and Beard [19], by Moreau [13] or, more recently, by Nedic and Ozdaglar [14], [15].

Network topologies change with time (as new nodes join and old nodes leave the network) or exhibit random behavior due to link failures, packet drops, node failure, etc. This motivated the investigation of consensus algorithms under a stochastic framework [6], [10], [11], [18], [20], [21]. In addition to network variability, nodes in sensor networks operate under limited computational, communication, and energy resources. These constraints have motivated the design of gossip algorithms, in which a node communicates with a randomly chosen neighbor. Studies of randomized gossip consensus algorithms can be found in [2], [22].

In this paper we introduce and analyze a generalized randomized gossip algorithm for achieving consensus. The algorithm acts on *convex metric spaces*, that are metric spaces endowed with a *convex structure*. We show that under the given algorithm, the agents' states converge to consensus with probability one and in the r^{th} mean sense. The convergence study is based on analyzing the dynamics of a set of stochastic differential equations driven by poisson counters. Additionally, for a particular network topology we investigate in more depth the rate of convergence of the first and second moment of the distances between the agents' states. We present instances of the generalized gossip algorithm for three convex

Ion Matei is with the Institute for Research in Electronics and Applied Physics at University of Maryland, College Park (email: imatei@umd.edu); Christoforos Somarakis and John S. Baras are with the Institute for Systems Research at University of Maryland, College Park (emails: csomarak@umd.edu, baras@umd.edu).

This material is based in part upon work supported by the NIST-ARRA Measurement Science and Engineering Fellowship Program award 70NANB10H026, through the University of Maryland, and in part upon work supported by the Army Research Office award number W911NF-08-1-0238 to Ohio State University.

metric spaces defined on the set of real numbers, the set of compact interval and the set of discrete random variables. The results of this paper, complements our previous results regarding the consensus problem on convex metric space, where only deterministic communication topologies are studied [8], [9].

The paper is organized as follows. Section II introduces the main concepts related to convex metric spaces. Section III formulates the problem and states our main results. Sections IV and V give the proof of our main results, together with pertinent preliminary results. In Section VI we present an in-depth analysis of the rate of convergence to consensus (in the first and second moments), for a particular network topology. Section VII shows instances of the generalized consensus algorithm for three convex metric spaces, defined on the sets of real numbers, compact intervals and discrete random variables, respectively.

Some basic notations: Given $W \in \mathbb{R}^{n\times n}$ by $[W]_{ij}$ we refer to the (i,j) element of the matrix. The *underlying graph* of W is a graph of order n without self loops, for which every edge corresponds to a *non-zero, off-diagonal* entry of W. We denote by $\chi_{\{A\}}$ the indicator function of the event A. Given two symmetric matrices M_1 and M_2, by $M_1 \succ M_2$ ($M_1 \succeq M_2$) we understand that $M_1 - M_2$ is a positive definite (semi-positive definite) matrix. Additionally, by $M_1 \prec M_2$ ($M_1 \preceq M_2$) we understand that $M_2 - M_1$ is a positive definite (semi-positive definite) matrix.

II. Convex Metric Spaces

In this section we introduce a set of definitions and basic results about convex metric spaces. Additional information about the following definitions and results can be found in [23],[24].

Definition 2.1 ([24], pp. 142): Let $(\mathcal{X},d)$ be a metric space. A mapping $\psi : \mathcal{X}\times\mathcal{X}\times[0,1] \to \mathcal{X}$ is said to be a *convex structure* on $\mathcal{X}$ if

$$d(u,\psi(x,y,\lambda)) \leq \lambda d(u,x) + (1-\lambda)d(u,y),\ \forall x,y,u \in \mathcal{X} \text{ and } \forall \lambda \in [0,1].$$

Definition 2.2 ([24], pp.142): The metric space $(\mathcal{X},d)$ together with the convex structure ψ is called a *convex metric space*, and is denoted henceforth by the triplet $(\mathcal{X},d,\psi)$.

Example 2.1: The most common convex metric space is defined on $\mathbb{R}^n$, together with the Euclidean distance and convex structure given by the standard convex combination operation. Indeed, for any $x,y,z \in \mathbb{R}^n$ and $\lambda \in [0,1]$, it follows that $\|z-(\lambda x+(1-\lambda)y)\| = \|\lambda(z-x)+(1-\lambda)(z-y)\| \leq \lambda\|z-x\|+(1-\lambda)\|z-y\|$, where the last inequality followed from the convexity of the norm operator.

Example 2.2 ([24]): Let $\mathcal{X}$ be the family of closed intervals, that is $\mathcal{X} = \{[a,b] \mid a \leq b,\ a,b \in \mathbb{R}\}$. For $I_i = [a_i,b_i]$, $I_j = [a_j,b_j]$ and $\lambda \in [0,1]$, we define a mapping ψ by $\psi(I_i,I_j,\lambda) = [\lambda a_i+(1-\lambda)a_j, \lambda b_i+(1-\lambda)b_j]$ and define a metric d in $\mathcal{X}$ by the Hausdorff distance, that is

$$d(I_i,I_j) = \max\{|a_i-a_j|,|b_i-b_j|\}.$$

Then $(\mathcal{X},d,\psi)$ is a convex metric space.

More examples can be found in [23] and [24]. In Section VII we introduce another interesting example of a convex metric space, defined on the set of discrete random variables taking values in a finite, countable set of real numbers.

Definition 2.3 ([24], pp. 144): A convex metric space $(\mathcal{X},d,\psi)$ is said to have *Property* (C) if every bounded decreasing net of nonempty closed convex subsets of $\mathcal{X}$ has a nonempty intersection.
Fortunately, convex matric spaces satisfying *Property* (C) are not that rare. Indeed, by Smulian's Theorem ([3], page 443), every weakly compact convex subset of a Banach space has *Property* (C).

The following definition introduces the notion of convex set in convex metric spaces.

Definition 2.4 ([24], pp. 143): Let $(\mathcal{X},d,\psi)$ be a convex metric space. A nonempty subset $K \subset \mathcal{X}$ is said to be *convex* if $\psi(x,y,\lambda) \in K$, $\forall x,y \in K$ and $\forall \lambda \in [0,1]$.

Let $\mathcal{P}(\mathcal{X})$ be the set of all subsets of $\mathcal{X}$. We define the set valued mapping $\tilde{\psi} : \mathcal{P}(\mathcal{X}) \to \mathcal{P}(\mathcal{X})$ as

$$\tilde{\psi}(A) \triangleq \{\psi(x,y,\lambda) \mid \forall x,y \in A, \forall \lambda \in [0,1]\},$$

where A is an arbitrary subset of $\mathcal{X}$.

In Proposition 1, pp. 143 of [24] it is shown that in a convex metric space, an arbitrary intersection of convex sets is also convex and therefore the next definition makes sense.

Definition 2.5 ([23], pp. 11): Let $(\mathcal{X}, d, \psi)$ be a convex metric space. The *convex hull* of the set $A \subset \mathcal{X}$ is the intersection of all convex sets in $\mathcal{X}$ containing A and is denoted by $co(A)$.

Another characterization of the convex hull of a set in $\mathcal{X}$ is given in what follows. By defining $A_m \triangleq \tilde{\psi}(A_{m-1})$ with $A_0 = A$ for some $A \subset \mathcal{X}$, it is discussed in [23] that the set sequence $\{A_m\}_{m\geq 0}$ is increasing and $\limsup_{m\to\infty} A_m$ exists, and $\limsup_{m\to\infty} A_m = \liminf_{m\to\infty} A_m = \lim_{m\to\infty} A_m = \bigcup_{m=0}^{\infty} A_m$.

Proposition 2.1 ([23], pp. 12): Let $(\mathcal{X}, d, \psi)$ be a convex metric space. The convex hull of a set $A \subset \mathcal{X}$ is given by

$$co(A) = \lim_{m\to\infty} A_m = \bigcup_{m=0}^{\infty} A_m.$$

It follows immediately from above that if $A_{m+1} = A_m$ for some m, then $co(A) = A_m$.

III. Problem formulation and main results

Let $(\mathcal{X}, d, \psi)$ be a convex metric space. We consider a set of n agents indexed by i, with states denoted by $x_i(t)$ taking values on $\mathcal{X}$, where t represents the continuous time.

A. Communication model

The communication among agents is subject to a communication network modeled by a directed graph $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the set of agents, and $E = \{(j, i) \mid j \text{ can send information to } i\}$ is the set of edges. In addition, we denote by $\mathcal{N}_i$ the inward neighborhood of agent i, i.e.,

$$\mathcal{N}_i \triangleq \{j \mid (j, i) \in E\},$$

where by assumption node i does not belong to the set $\mathcal{N}_i$. We make the following connectivity assumption.

Assumption 3.1: The graph $G = (V, E)$ is strongly connected.

B. Randomized gossip algorithm

In the following we describe the mechanism used by the agents to update their states. Agents can be in two modes: *sleep* mode and *update* mode. Let $N_i(t)$ be a Poisson counter associated to agent i. In the sleep mode, the agents maintain their states unchanged. An agent i exits the sleep mode and enters the update mode, when the associated counter $N_i(t)$ increments its value. Let t_i be a time-instant at which the Poisson counter $N_i(t)$ increments its value. Then at t_i, agent i picks agent j with probability $p_{i,j}$, where $j \in \mathcal{N}_i$ and updates its state according to the rule

$$x_i(t_i^+) = \psi(x_i(t_i), x_j(t_i), \lambda_i), \tag{1}$$

where $\lambda_i \in [0,\ 1)$, ψ is the convex structure and $\sum_{j\in\mathcal{N}_i} p_{i,j} = 1$. By $x_i(t_i^+)$ we understand the value of $x_i(t)$ immediately after the instant update at time t_i, which can be also written as

$$x_i(t_i^+) = \lim_{t\to t_i,\ t>t_i} x_i(t),$$

which implies that $x_i(t)$ is a left-continuous function of t. After agent i updates its state according to the above rule, it immediately returns to the sleep mode, until the next increase in value of the counter $N_i(t)$.

Assumption 3.2: The Poisson counters $N_i(t)$ are independent and with rate μ_i, for all i.

A similar form of the above algorithm (the Poisson counters are assumed to have the same rates) was extensively studied in [2], in the case where $\mathcal{X} = \mathbb{R}$.

We first note that since the agents update their state at random times, the distances between agents are random processes. Let $d(x_i(t), x_j(t))$ be the distance between the states of agents i and j, at time t. We introduce the following convergence definitions.

Definition 3.1: We say that the agents converge to consensus *with probability one* if

$$Pr\left(\lim_{t\to\infty} \max_{i,j} d(x_i(t), x_j(t)) = 0\right) = 1.$$

Definition 3.2: We say that the agents converge to consensus in r^{th} *mean sense* if

$$\lim_{t\to\infty} E\left[d(x_i(t), x_j(t))^r\right] = 0, \forall\ (i, j),\ i \neq j.$$

The following theorems state our main convergence results.

Theorem 3.1: Under Assumptions 3.1 and 3.2 and under the randomized gossip algorithm, the agents converge to consensus in r^{th} mean, in the sense of Definition 3.2.

Theorem 3.2: Under Assumptions 3.1 and 3.2 and under the randomized gossip algorithm, the agents converge to consensus with probability one, in the sense of Definition 3.1.

The above results show that the distances between the agents' states converge to zero. The following Corollary shows that in fact, for convex metric spaces satisfying *Property* (C), the states of the agents converge to some point in the convex metric space.

Corollary 3.1: Under Assumptions 3.1 and 3.2 and under the randomized gossip algorithm operating on convex metric spaces satisfying *Property* (C), for any sample path ω of state processes, there exits $x^* \in \mathcal{X}$ (that depends on ω and the initial conditions $x_i(0)$) such that

$$\lim_{t\to\infty} d(x_i(t,\omega), x^*(\omega)) = 0.$$

In other words, the states of the agents converge to some point of the convex metric space with probability one.

IV. Preliminary results

In this section we construct the stochastic dynamics of the vector of distances between agents. Let t_i be a time-instant at which counter $N_i(t)$ increments its value. Then according to the gossip algorithm, at time t_i^+ the distance between agents i and j is given by

$$d(x_i(t_i^+), x_j(t_i^+)) = d(\psi(x_i(t_i), x_l(t_i), \lambda_i), x_j(t_i)), \text{ with probability } p_{i,l}. \tag{2}$$

Let $\theta_i(t)$ be an independent and identically distributed (i.i.d.) random process, such that $Pr(\theta_i(t) = l) = p_{i,l}$ for all $l \in \mathcal{N}_i$ and for all t. It follows that (2) can be equivalently written as

$$d(x_i(t_i^+), x_j(t_i^+)) = \sum_{l\in\mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} d(\psi(x_i(t_i), x_l(t_i), \lambda_i), x_j(t_i)), \tag{3}$$

where $\chi_{\{\cdot\}}$ denotes the indicator function. Using the inequality property of the convex structure introduced in Definition 2.1, we further get

$$d(x_i(t_i^+), x_j(t_i^+)) \leq \lambda_i d(x_i(t_i), x_j(t_i)) + (1-\lambda_i) \sum_{l\in\mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} d(x_l(t_i), x_j(t_i)). \tag{4}$$

Assuming that t_j is a time-instant at which the Poisson counter $N_j(t)$ increments its value, in a similar manner as above we get that

$$d(x_i(t_j^+), x_j(t_j^+)) \leq \lambda_j d(x_i(t_j), x_j(t_j)) + (1-\lambda_j) \sum_{l\in\mathcal{N}_j} \chi_{\{\theta_j(t_j)=l\}} d(x_l(t_j), x_i(t_j)). \tag{5}$$

Consider now the scalars $\eta_{i,j}(t)$ which follow the same dynamics as the distance between agents i and j, but with equality, that is,

$$\eta_{i,j}(t_i^+) = \lambda_i \eta_{i,j}(t_i) + (1-\lambda_i) \sum_{l \in \mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} \eta_{j,l}(t_i), \tag{6}$$

and

$$\eta_{i,j}(t_j^+) = \lambda_j \eta_{i,j}(t_j) + (1-\lambda_j) \sum_{l \in \mathcal{N}_j} \chi_{\{\theta_j(t_j)=l\}} \eta_{i,l}(t_j), \tag{7}$$

with $\eta_{i,j}(0) = d(x_i(0), x_j(0))$.

Remark 4.1: Note that the index pair of η refers to the distance between two agents i and j. As a consequence $\eta_{i,j}$ and $\eta_{j,i}$ will be considered the same objects, and counted only once.

Proposition 4.1: The following inequalities are satisfied with probability one:

$$\eta_{i,j}(t) \geq 0, \tag{8}$$

$$\eta_{i,j}(t) \leq \max_{i,j} \eta_{i,j}(0), \tag{9}$$

$$d(x_i(t), x_j(t)) \leq \eta_{i,j}(t), \tag{10}$$

for all $i \neq j$ and $t \geq 0$.

Proof: Inequalities (8) and (9) follow immediately, noting that for any sample path of the Poisson counters, $\eta_{i,j}(t)$ are updated by performing convex combinations of non-negative quantities. To show inequality (10) we can use an inductive argument. Let t_i be the time instant at which the counter $N_i(t)$ increments its value and assume that $d(x_i(t_i), x_j(t_i)) \leq \eta_{i,j}(t_i)$ for all i, j. Immediately after t_i, the new value of $d(x_i(t), x_j(t))$ is given by

$$d(x_i(t_i^+), x_j(t_i^+)) \leq \lambda_i d(x_i(t_i), x_j(t_i)) + (1-\lambda_i) \sum_{l \in \mathcal{N}_j} \chi_{\{\theta_j(t_i)=l\}} d(x_l(t_i), x_i(t_i)) \leq$$

$$\leq \lambda_i \eta_{i,j}(t_i) + (1-\lambda_i) \sum_{l \in \mathcal{N}_i} \chi_{\{\theta_i(t_i)=l\}} \eta_{j,l}(t_i) = \eta_{i,j}(t_i^+).$$

Therefore after each increment of counter $N_i(t)$, we get that

$$d(x_i(t_i^+), x_j(t_i^+)) \leq \eta_{i,j}(t_i^+).$$

Using the same argument for all Poison counters, inequality (10) follows. ■

We now construct the stochastic differential equation satisfied by $\eta_{i,j}(t)$. From equations (6) and (7) we note that $\eta_{i,j}(t)$ at time t_i and t_j satisfies the solution of a stochastic differential equation driven by Poisson counters. Namely, we have

$$d\eta_{i,j}(t) = \left[-(1-\lambda_i)\eta_{i,j}(t) + (1-\lambda_i) \sum_{l \in \mathcal{N}_i} \chi_{\{\theta_i(t)=l\}} \eta_{j,l}(t) \right] dN_i(t) +$$

$$\left[-(1-\lambda_j)\eta_{i,j}(t) + (1-\lambda_j) \sum_{m \in \mathcal{N}_j} \chi_{\{\theta_j(t)=m\}} \eta_{i,m}(t) \right] dN_j(t). \tag{11}$$

Let us now define the $\bar{n}$ dimensional vector $\boldsymbol{\eta} = (\eta_{i,j})$, where $\bar{n} = \frac{n(n-1)}{2}$ (since (i, j) and (j, i) correspond to the same distance variable). Equation (11) can be compactly written as

$$d\boldsymbol{\eta}(t) = \sum_{(i,j), i \neq j} \Phi_{i,j}(\theta_i(t)) \boldsymbol{\eta}(t) dN_i(t) + \sum_{(i,j), i \neq j}, \Psi_{i,j}(\theta_j(t)) \boldsymbol{\eta}(t) dN_j(t). \tag{12}$$

where the $\bar{n}\times\bar{n}$ dimensional matrices $\Phi_{i,j}(\theta_i(t))$ and $\Psi_{i,j}(\theta_j(t))$ are defined as:

$$\Phi_{i,j}(\theta_i(t)) = \begin{cases} -(1-\lambda_i) & \text{at entry } [(i,j)(i,j)] \\ (1-\lambda_i)\chi_{\{\theta_i(t)=l\}} & \text{at entries } [(i,j)(l,j)], \quad l \in \mathcal{N}_i,\ l \neq j, l \neq i \\ 0 & \text{all other entries ,} \end{cases} \tag{13}$$

and

$$\Psi_{i,j}(\theta_j(t)) = \begin{cases} -(1-\lambda_j) & \text{at entry } [(i,j)(i,j)] \\ (1-\lambda_j)\chi_{\{\theta_j(t)=m\}} & \text{at entries } [(i,j)(m,i)], \quad m \in \mathcal{N}_j, m \neq j, m \neq i \\ 0 & \text{all other entries .} \end{cases} \tag{14}$$

The dynamics of the first moment of the vector $\boldsymbol{\eta}(t)$ is given by

$$\frac{d}{dt}E\{\boldsymbol{\eta}(t)\} = \sum_{(i,j),i\neq j} E\left\{\Phi_{i,j}(\theta_i(t))\boldsymbol{\eta}(t)\mu_i + \Psi_{i,j}(\theta_j(t))\boldsymbol{\eta}(t)\mu_j\right\}. \tag{15}$$

Using the independence of the random processes $\theta_i(t)$, we can further write

$$\frac{d}{dt}E\{\boldsymbol{\eta}(t)\} = \boldsymbol{W}E\{\boldsymbol{\eta}(t)\}, \tag{16}$$

where $\boldsymbol{W}$ is a $\bar{n}\times\bar{n}$ dimensional matrix whose entries are given by

$$[\boldsymbol{W}]_{(i,j),(l,m)} = \begin{cases} -(1-\lambda_i)\mu_i - (1-\lambda_j)\mu_j & l = i \text{ and } m = j \\ (1-\lambda_i)\mu_i p_{i,l} & l \in \mathcal{N}_i,\ m = j,\ l \neq j, \\ (1-\lambda_j)\mu_j p_{j,m} & l = i,\ m \in \mathcal{N}_j,\ m \neq i, \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

The following Lemma studies the properties of matrix $\boldsymbol{W}$, introduced above.

Lemma 4.1: Let $\boldsymbol{W}$ be the $\bar{n}\times\bar{n}$ dimensional matrix defined in (17). Under Assumption 3.1, the following properties hold:

(a) Let $\bar{G}$ be the directed graph (without self loops) corresponding to the matrix $\boldsymbol{W}$, that is, a link from (l,m) to (i,j) exists in $\bar{G}$ if $[\boldsymbol{W}]_{(i,j),(l,m)} > 0$. Then $\bar{G}$ is strongly connected.

(b) The row sums of matrix $\boldsymbol{W}$ are non-positive, i.e.,

$$\sum_{(l,m),l\neq m} [\boldsymbol{W}]_{(i,j)(l,m)} \leq 0,\ \forall (i,j), i \neq j.$$

(c) There exits at least one row (i^*, j^*) of $\boldsymbol{W}$ whose sum is negative, that is,

$$\sum_{(l,m),l\neq m} [\boldsymbol{W}]_{(i^*,j^*)(l,m)} < 0.$$

Proof: (a) Consider the paths from agent i to agent l ($i = a_0, \ldots, a_{k-1} = l$) and from agent j to agent m ($j = b_0, \ldots, b_{k-1} = m$). By Assumption 3.1 one can always construct same length paths from any agent to another since they belong to the same equivalence class. We use (17) and dictate and algorithmic proof which ensures the existence of a path from (i,j) to (l,m) on $\bar{G}$:

```
for z = 0 to k−1
  If a_{z+1} ≠ b_z
     jump to (a_{z+1}, b_z)
  else if b_{z+1} ≠ a_z
     jump to (a_z, b_{z+1})
  else
     jump to (a_{z+1}, b_{z+1})
```

Note that the last case, the condition essentially means that the walker does not move. (We consider the nodes (i,j) and (j,i) as identical). For arbitrary (i,j) and (l,m) the result follows.

(b) Consider a row (i,j). For convenience, let us define the following positive scalars

$$\xi_i \triangleq (1-\lambda_i)\mu_i \text{ and } \xi_j \triangleq (1-\lambda_j)\mu_j. \tag{18}$$

We can express the sum of row (i,j)'s entries as

$$\sum_{(l,m)}[\boldsymbol{W}]_{(i,j)(l,m)} = -(\xi_i+\xi_j)+\xi_i \sum_{l\in\mathcal{N}_i, l\neq j, m=j} p_{i,l} + \xi_j \sum_{m\in\mathcal{N}_j, m\neq i, l=i} p_{j,m} \leq$$

$$\leq -(\xi_i+\xi_j)+\xi_i+\xi_j = 0.$$

(c) Consider an arbitrary row (i,j). The row (i,j) would sum up to zero in two cases. In the first case, $i \notin \mathcal{N}_j$ and $j \notin \mathcal{N}_i$, which imply

$$\sum_{l\in\mathcal{N}_i, l\neq j, m=j} p_{i,l} = 1 \text{ and } \sum_{m\in\mathcal{N}_j, m\neq i, l=i} p_{j,m} = 1,$$

and therefore

$$\sum_{(l,m)}[\boldsymbol{W}]_{(i,j)(l,m)} = -(\xi_i+\xi_j)+\xi_i+\xi_j = 0.$$

However, having $i \notin \mathcal{N}_j$ and $j \notin \mathcal{N}_i$ for all i and j would imply the communication graph $G = (V,E)$ to be disconnected, contradicting Assumption 3.1. In the second case, $i \in \mathcal{N}_j$ and $j \in \mathcal{N}_i$ and $|\mathcal{N}_i| = 1$ and $|\mathcal{N}_j| = 1$ (that is, node i has only one neighbor, namely j and j has only one neighbor, namely i). In this case

$$\sum_{l\in\mathcal{N}_i, l\neq j, m=j} p_{i,l} = p_{i,j} = 1 \text{ and } \sum_{m\in\mathcal{N}_j, m\neq i, l=i} p_{j,m} = p_{j,i} = 1,$$

and consequently

$$\sum_{(l,m)}[\boldsymbol{W}]_{(i,j)(l,m)} = -(\xi_i+\xi_j)+\xi_i+\xi_j = 0.$$

But this case would imply that the nodes i and j are separated from all other nodes in the graph $G = (V,E)$, which would contradict the connectivity Assumption 3.1. Therefore, there must exist at least one row (i^*, j^*) such that

$$\sum_{(l,m)}[\boldsymbol{W}]_{(i^*,j^*)(l,m)} < 0.$$

$\blacksquare$

Consider now the matrix $\boldsymbol{Q} \triangleq I + \epsilon\boldsymbol{W}$, where I is the identity matrix and ϵ is a positive scalar satisfying the strict inequality

$$0 < \epsilon < \frac{1}{\max_{i,j}(\xi_i+\xi_j)}, \tag{19}$$

where ξ_i and ξ_j were defined in (18).

The following Corollary follows from the previous Lemma and describes the properties of matrix $\boldsymbol{Q}$.

Corollary 4.1: Matrix $\boldsymbol{Q}$ has the following properties:

(a) The directed graph (without self loops) corresponding to matrix $\boldsymbol{Q}$ (that is, a link from (l,m) to (i,j) exists if $[\boldsymbol{Q}]_{(i,j),(l,m)} > 0$) is strongly connected.

(b) $\boldsymbol{Q}$ is a non-negative matrix with positive diagonal elements.

(c) The rows of $\boldsymbol{Q}$ sum up to a positive value not larger than one, that is,

$$\sum_{(l,m), l\neq m}[\boldsymbol{Q}]_{(i,j)(l,m)} \leq 1, \ \forall (i,j).$$

(d) There exits at least one row (i^*, j^*) of $\boldsymbol{Q}$ which sums up to a positive value strictly smaller than one, that is,

$$\sum_{(l,m),l\neq m}[\boldsymbol{Q}]_{(i^*,j^*)(l,m)} < 1.$$

Proof: Noting that the directed graph (without self loops) corresponding to matrix $\boldsymbol{Q}$ is identical to the one corresponding to matrix $\boldsymbol{W}$, part (a) follows. The diagonal elements of $\boldsymbol{Q}$ are given by

$$[\boldsymbol{Q}]_{(i,j)(i,j)} = 1-\epsilon(\xi_i+\xi_j).$$

Using the fact that $0<\epsilon(\xi_i+\xi_j)<1$, and the obvious observation that the non-diagonal elements are non-negative, we obtain part (b). The sum of row (i,j) entries is given by

$$\sum_{(l,m)}[\boldsymbol{Q}]_{(i,j)(l,m)} = 1+\epsilon\sum_{(l,m)}[\boldsymbol{W}]_{(i,j)(l,m)},$$

and using parts (b) and (c) of Lemma 4.1, part (c) and (d) of the current Corollary follow, respectively. ■

Remark 4.2: The above Corollary says that the matrix $\boldsymbol{Q}$ is an *irreducible*, *substochastic* matrix. In addition, choosing $\gamma \geq \max_{i,j}\frac{1}{[\boldsymbol{Q}]_{(i,j)(i,j)}}$, it follows that we can find a non-negative, irreducible matrix $\tilde{\boldsymbol{Q}}$ such that $\gamma\boldsymbol{Q} = I+\tilde{\boldsymbol{Q}}$. Using a result on converting non-negativity and irreducibility to positivity ([12], page 672), we get that $(I+\tilde{\boldsymbol{Q}})^{\bar{n}-1} = \gamma^{\bar{n}-1}\boldsymbol{Q}^{\bar{n}-1} > 0$, and therefore $\boldsymbol{Q}$ is a *primitive* matrix. The existence of γ is guaranteed by the fact that $\boldsymbol{Q}$ has positive diagonal entries.

We have the following result on the spectral radius of $\boldsymbol{Q}$, denoted by $\rho(\boldsymbol{Q})$.

Lemma 4.2: The spectral radius of matrix $\boldsymbol{Q}$ is smaller than one, that is,

$$\rho(\boldsymbol{Q}) < 1.$$

Proof: Let $(\bar{i},\bar{j})$ denote an entry of matrix $\boldsymbol{Q}$. As mentioned in Remark 4.2, $\boldsymbol{Q}$ is a primitive matrix, and therefore there exits a k (which in our case is $\bar{n}-1$), such that $\boldsymbol{Q}^k$ has all entries positive, that is

$$[\boldsymbol{Q}^k]_{\bar{i},\bar{j}} > 0,\ \forall \bar{i},\bar{j}.$$

In addition, since $\boldsymbol{Q}$ is a substochastic matrix we have that

$$\sum_{\bar{j}=1}^{\bar{h}}[\boldsymbol{Q}^k]_{\bar{i},\bar{j}} \leq 1,\ \forall \bar{i}.$$

By part (d) of Lemma 4.2, we have that there exits a $\bar{j}^*$ such that

$$\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j}^*,\bar{h}} < 1.$$

Let us not consider the matrix $\boldsymbol{Q}^{k+1} = \boldsymbol{Q}^k Q$. The sum of the $\bar{i}^{th}$ row entries of matrix $\boldsymbol{Q}^{k+1}$ is given by

$$\sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k+1}]_{\bar{i},\bar{j}} = \sum_{\bar{j}=1,\bar{j}\neq\bar{j}^*}^{\bar{n}}[\boldsymbol{Q}^k]_{\bar{i},\bar{j}}\left(\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j},\bar{h}}\right)+[\boldsymbol{Q}^k]_{\bar{i},\bar{j}^*}\left(\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j}^*,\bar{h}}\right).$$

But since

$$0\leq\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j},\bar{h}} \leq 1,\ \forall \bar{j},$$

it follows that

$$\sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k+1}]_{\bar{i},\bar{j}} \leq \sum_{\bar{j}=1,\bar{j}\neq\bar{j}^*}^{\bar{n}}[\boldsymbol{Q}^k]_{\bar{i},\bar{j}}+[\boldsymbol{Q}^k]_{\bar{i},\bar{j}^*}\left(\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j}^*,\bar{h}}\right),$$

or

$$\sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k+1}]_{\bar{i},\bar{j}} \leq \sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k}]_{\bar{i},\bar{j}} - [\boldsymbol{Q}^{k}]_{\bar{i},\bar{j}^*}\left(1-\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j}^*,\bar{h}}\right).$$

Additionally, since $0 \leq \sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k}]_{\bar{i},\bar{j}} \leq 1$, $[\boldsymbol{Q}^{k}]_{\bar{i},\bar{j}^*} > 0$, $\sum_{\bar{h}=1}^{\bar{n}}[\boldsymbol{Q}]_{\bar{j}^*,\bar{h}} < 1$, and $\sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k+1}]_{\bar{i},\bar{j}} \geq 0$, it must be that

$$\sum_{\bar{j}=1}^{\bar{n}}[\boldsymbol{Q}^{k+1}]_{\bar{i},\bar{j}} < 1, \ \forall \bar{i}.$$

But this means that the infinity norm of $\boldsymbol{Q}^{k+1}$ is smaller than one, that is

$$||\boldsymbol{Q}^{k+1}||_\infty < 1,$$

or that $\rho(\boldsymbol{Q}^{k+1}) < 1$, which in turns implies that

$$\rho(\boldsymbol{Q}) < 1.$$

By the Peron-Frobenius Theorem for non-negative, irreducible matrices ([12], page 673) $\rho(\boldsymbol{Q})$ corresponds to the positive eigenvalue of $\boldsymbol{Q}$, larger than the absolute values of all other eigenvalues. ■

V. Proof of the main results

In this section we prove our main results introduced in Section III.

A. Proof of Theorem 3.1

We first show that the vector $\boldsymbol{\eta}(t)$ converges to zero in mean. By Lemma 4.2 we have that the spectral radius of $\boldsymbol{Q}$ is smaller than one, that is

$$\rho(\boldsymbol{Q}) < 1,$$

where $\rho(\boldsymbol{Q}) = \max_{\bar{i}} |\lambda_{\bar{i},\boldsymbol{Q}}|$, with $\lambda_{\bar{i},\boldsymbol{Q}}$, $\bar{i} = 1, \ldots, \bar{n}$ being the eigenvalues of $\boldsymbol{Q}$. This also means that

$$Re(\lambda_{\bar{i},\boldsymbol{Q}}) < 1, \ \forall \bar{i}. \tag{20}$$

But since $\boldsymbol{W} = \frac{1}{\epsilon}(\boldsymbol{Q} - I)$, it follows that the real part of the eigenvalues of $\boldsymbol{W}$ are given by

$$Re(\lambda_{\bar{i},\boldsymbol{W}}) = \frac{1}{\epsilon}\left(Re(\lambda_{\bar{i},\boldsymbol{Q}}) - 1\right) < 0, \ \forall \bar{i},$$

where the last inequality follows from (20). Therefore, the dynamics

$$\frac{d}{dt}E\{\boldsymbol{\eta}(\boldsymbol{t})\} = \boldsymbol{W}E\{\boldsymbol{\eta}(t)\}$$

is asymptotically stable, and hence $\boldsymbol{\eta}(t)$ converges in mean to zero.

We now show that $\boldsymbol{\eta}(t)$ converges in r^{th} mean, for any $r \geq 1$. We showed above that $\eta_{i,j}(t)$ converges in mean to zero, for any $i \neq j$. But this also implies that $\eta_{i,j}(t)$ converges to zero in probability (Theorem 3, page 310, [5]), and therefore, for any $\delta > 0$

$$\lim_{t\to\infty} Pr(\eta_{i,j}(t) > \delta) = 0. \tag{21}$$

Using the indicator function, the quantity $\eta_{i,j}(t)$ can be expressed as

$$\eta_{i,j}(t) = \eta_{i,j}(t)\chi_{\{\eta_{i,j}(t)\leq\delta\}} + \eta_{i,j}(t)\chi_{\{\eta_{i,j}(t)>\delta\}},$$

for any $\delta > 0$. Using (9) of Proposition 4.1 we can further write

$$\eta_{i,j}(t) \leq \delta\chi_{\{\eta_{i,j}(t)\leq\delta\}} + \max_{i,j}\eta_{i,j}(0)\chi_{\{\eta_{i,j}(t)>\delta\}},$$

or

$$\eta_{i,j}(t)^r \leq \delta^r \chi_{\{\eta_{i,j}(t)\leq\delta\}} + \left(\max_{i,j} \eta_{i,j}(0)\right)^r \chi_{\{\eta_{i,j}(t)>\delta\}},$$

where to obtain the previous inequality we used the fact that $\chi_{\{\eta_{i,j}(t)\leq\delta\}}\chi_{\{\eta_{i,j}(t)>\delta\}} = 0$. Using the expectation operator, we obtain

$$E\{\eta_{i,j}(t)^r\} \leq \delta^r \Pr(\eta_{i,j}(t) \leq \delta) + \left(\max_{i,j} \eta_{i,j}(0)\right)^r Pr(\eta_{i,j}(t) > \delta).$$

Taking t to infinity results in

$$\limsup_{t\to\infty} E\{\eta_{i,j}(t)^r\} \leq \delta^r, \ \forall\delta > 0,$$

and since δ can be made arbitrarily small, we have that

$$\lim_{t\to\infty} E\{\eta_{i,j}(t)^r\} = 0, \ \forall r \geq 1.$$

Using (10) of Proposition 4.1, the result follows.

B. Proof of Theorem 3.2

In the following we show that $\boldsymbol{\eta}(t)$ converge to zero almost surely. Equations (6) and (7) show that with probability one $\eta_{i,j}(t)$ is non-negative and that for any $t_2 \leq t_1$, with probability one $\eta_{i,j}(t_2)$ belongs to the convex hull generated by $\{\eta_{l,m}(t_1) \mid \text{for all pairs } (l,m)\}$. But this also implies that with probability one

$$\max_{i,j} \eta_{i,j}(t_2) \leq \max_{i,j} \eta_{i,j}(t_1). \tag{22}$$

Hence for any sample path of the random process $\boldsymbol{\eta}(t)$, the sequence $\{\max_{i,j} \eta_{i,j}(t)\}_{t\geq 0}$ is monotone decreasing and lower bounded. Using the monotone convergence theorem, we have that for any sample path ω, there exits $\tilde{\eta}(\omega)$ such that

$$\lim_{t\to\infty} \max_{i,j} \eta_{i,j}(t,\omega) = \tilde{\eta}(\omega),$$

or similarly

$$Pr\left(\lim_{t\to\infty} \max_{i,j} \eta_{i,j}(t) = \tilde{\eta}\right) = 1.$$

In the following we show that $\tilde{\eta}$ must be zero with probability one. We achieve this by showing that there exits a subsequence of $\{\max_{i,j} \eta_{i,j}(t)\}_{t\geq 0}$ that converges to zero with probability one.

In Theorem 3.1 we proved that $\boldsymbol{\eta}(t)$ converges to zero in the r^{th} mean. Therefore, for any pair (i, j) and (l,m) we have that $E\{\eta_{i,j}(t)\eta_{l,m}(t)\}$ converge to zero. Moreover, since

$$E\{\eta_{i,j}(t)\eta_{l,m}(t)\} \leq \max_{i,j} \eta_{i,j}(0) E\{\eta_{l,m}(t)\},$$

and since $E\{\eta_{l,m}(t)\}$ converges to zero exponentially, we have that $E\{\eta_{i,j}(t)\eta_{l,m}(t)\}$ converges to zero exponentially as well.

Let $\{t_k\}_{k\geq 0}$ be a time sequence such that $t_k = kh$, for some $h > 0$. From above, it follows that $E\{\|\boldsymbol{\eta}(t_k)\|^2\}$ converges to zero geometrically. But this is enough to show that the sequence $\{\boldsymbol{\eta}(t_k)\}_{k\geq 0}$ converges to zero, with probability one by using the Borel-Cantelli lemma (Theorem 10, page 320, [5]). Therefore, $\tilde{\eta}$ must be zero. Using (9) of Proposition 4.1, the result follows.

C. Proof of Corollary 3.1

The main idea of the proof consists of showing that the convex hull of the states of the agents converge to one point, for any sample path of the states processes. Let ω be a sample path of the state process and let $\{t_k\}_{k\geq 0}$ be the time instants at which the Poisson counters increase their values, corresponding to this sample path. Additionally, let A_k be the set of the agents' states at time t_k, that is $A_k = \left\{x_j(t_k), j = 1\dots n\right\}$. According to Definition 2.5, Proposition 2.1 and equation (1) of the randomized gossip algorithm, we have that

$$x_i(t_{k+1}) \in co(A_k), \ \forall i.$$

But this also implies the next convex hull's inclusion

$$co(A_{k+1}) \subseteq co(A_k).$$

From the theory of limit of sequence of sets it follows that there exits a set A_∞ such that

$$\limsup co(A_k) = \liminf co(A_k) = \lim co(A_k) = A_\infty,$$

where $A_\infty = \bigcap_{k\geq 0} co(A_k)$.

Denoting the diameter of the set A_k by

$$diam(A_k) = \sup\{d(x,y) \mid x,y \in A_k\},$$

from Proposition 2 of [23], we have that

$$diam(A_k) = diam(co(A_k)).$$

Additionally, in Theorem 3.2 we showed that

$$\lim_{t\to\infty} d(x_i(t), x_j(t)) = 0, \ \forall \ (i,j),$$

with probability one and therefore, the same is true for the sample path ω, that is

$$\lim_{k\to\infty} d(x_i(t_k), x_j(t_k)) = 0, \ \forall \ (i,j).$$

But this means that

$$\lim_{k\to\infty} diam(A_k) = \lim_{k\to\infty} diam(co(A_k)) = 0,$$

and therefore $diam(A_\infty) = 0$. But since the convex metric space on which the randomized gossip algorithm operates satisfies *Property* (C), and the sets A_k are bounded and closed, it follows that the set A_∞ is non-empty. Consequently, there exits a point x^*, which may depend on ω, such that $A_\infty = x^*$, and the result follows.

VI. The rate of convergence of the generalized gossip consensus algorithm under complete and uniform connectivity

We note that under our general problem setup it is difficult to get explicit formulas for the rate of convergence to consensus, in the first and second moments. We are able however to obtain explicit results for the aforementioned rates of convergence under specific assumptions on the topology of the graph, on the parameters of the Poisson counters and on the convex structure.

Assumption 6.1: The Poisson counters have the same rate, that is $\mu_i = \mu$ for all i. Additionally, the parameters used by the agents in the convex structure are equal, that is $\lambda_i = \lambda$, for all i. In the update mode, each agent i picks one of the rest $n-1$ agents uniformly, that is $\mathcal{N}_i = \mathcal{N} - \{i\}$ and $p_{i,j} = \frac{1}{n-1}$, for all $j \in \mathcal{N}_i$.

The following two Propositions give upper bounds on the rate of convergence for the first and second moments of the distance between agents, under Assumption 6.1.

Proposition 6.1: Under Assumptions 3.1, 3.2 and 6.1, the first moment of the distances between agents' states, using the generalized gossip algorithm converges exponentially to zero, that is

$$E\{d(x_i(t),x_j(t))\} \leq c_1 e^{\alpha_1 t}, \text{ for all pairs } (i,j), \tag{23}$$

where $\alpha_1 = -\frac{2(1-\lambda)\mu}{n-1}$ and c_1 is a positive scalar depending of the initial conditions.

Proof: By Proposition 4.1, with probability one we have that for any pair (i,j) $d(x_i(t),x_j(t)) \leq \eta_{i,j}(t)$ and therefore $E\{d(x_i(t),x_j(t))\} \leq E\{\eta_{i,j}(t)\}$. But the convergence of $E\{\eta_{i,j}(t)\}$ is determined by equation (16) and in particular by the eigenvalues of matrix $\boldsymbol{W}$, which are studied in what follows. From (17) it immediately follows that $\boldsymbol{W}$ is a symmetric matrix and that every diagonal element is $-2(1-\lambda)\mu$. Note the enumeration of the vertex set of graph $\bar{\mathcal{G}}$ as (i,j) with $(i<j)$. Consider an arbitrary node (i,j) and write the element of the corresponding row in the following convenient form

$$\begin{array}{l}
(1,2),\ (1,3),\ \ldots,\ (1,n)\\
(2,3),\ (2,4),\ \ldots,\ (2,n)\\
\ldots\\
(i-1,i),\ (i-1,i+1),\ \ldots,\ (i-1,n)\\
\hline
(i,i+1),\ \ldots,\ (i,n)\\
\ldots\\
(j-1,j),\ \ldots,\ (j-1,n)\\
\hline
(j,j+1),\ \ldots,\ (j,n)\\
\hline
(j+1,j+2),\ \ldots,\ (n-1,n)
\end{array}$$

where we split it with horizontal lines in 5 segments (numbered 1 through 5 from top to bottom). Following (17) observe that excluding the diagonal, the matrix has exactly $2i-2$ positive elements in segment 1, $n-i-1$ positive elements in segment 2, $j-i-1$ positive elements in segment 3, $n-j$ positive elements in segment 4 and 0 positive elements in segment 5. Therefore the total number of off-diagonal entries in a row is $2n-4$. Again, (17) dictates that the value in any positive element is $\mu\frac{1-\lambda}{n-1}$. As a consequence, we conclude that the sum of every row is $\alpha_1 = -\frac{2(1-\lambda)\mu}{n-1}$, that is obviously the eigenvalue of the right eigenvector $\mathbb{1}_{\bar{n}}$, that is the vector of all ones. Noting that $\boldsymbol{W}$ is symmetric all eigenvalues are real and by Greshgorin' theorem (Theorem 7.2.1, page 320, [4]) they must lie in the circle $(-2(1-\lambda)\mu, r)$ where $r = 2(1-\lambda)\mu\frac{n-2}{n-1}$ is the sum of the non zero, off-diagonal elements of the rows. Note that the eigenvalue α_1 lies exactly on the boundary of the circle, in the negative half plane. This leads us to conclude that this is indeed the maximum one. Therefore, there exits a positive scalar c_1 which depends on the initial conditions such that

$$E\{\eta_{i,j}(t)\} \leq c_1 e^{\alpha_1 t}, \quad \text{for all } (i,j),$$

from where the result follows. ■

Proposition 6.2: Under Assumptions 3.1, 3.2 and 6.1, the second moment of the distances between agents' states, using the generalized gossip algorithm converges exponentially to zero, that is

$$E\{d(x_i(t),x_j(t))^2\} \leq c_2 e^{\alpha_2 t}, \text{ for all pair } (i,j),$$

where $\alpha_2 = -\mu\frac{2(1-\lambda^2)}{n-1}$ and c_2 is a positive scalar depending of the initial distances between agents.

Proof: As before, by Proposition 4.1, with probability one we have that for any pair (i,j) $d(x_i(t),x_j(t)) \leq \eta_{i,j}(t)$ and therefore $E\{d(x_i(t),x_j(t))^2\} \leq E\{\eta_{i,j}(t)^2\}$. But $E\{\eta_{i,j}(t)^2\} \leq E\{\|\boldsymbol{\eta}(t)\|^2\}$, for any pair (i,j) and therefore is sufficient to study the convergence properties of the right-hand side of the previous inequality. Using Ito's rule we can differentiate the quantity $\|\boldsymbol{\eta}(t)\|^2$ and obtain

$$\frac{d}{dt}\|\boldsymbol{\eta}(t)\|^2 = \sum_{i,j}\boldsymbol{\eta}(t)'\left[\Phi_{i,j}(\theta_i(t)) + \Phi_{i,j}(\theta_i(t))' + \Phi_{i,j}(\theta_i(t))'\Phi_{i,j}(\theta_i(t))\right]\boldsymbol{\eta}(t)dN_i(t)+$$

$$+\sum_{i,j}\boldsymbol{\eta}(t)'\left[\Psi_{i,j}(\theta_j(t))+\Psi_{i,j}(\theta_j(t))'+\Psi_{i,j}(\theta_j(t))'\Psi_{i,j}(\theta_j(t))\right]\boldsymbol{\eta}(t)dN_j(t),$$

from where we get

$$\frac{d}{dt}E\{\|\boldsymbol{\eta}(t)\|^2\}=\sum_{i,j}E\{\boldsymbol{\eta}(t)'\left[\Phi_{i,j}(\theta_i(t))+\Phi_{i,j}(\theta_i(t))'+\Phi_{i,j}(\theta_i(t))'\Phi_{i,j}(\theta_i(t))\right]\boldsymbol{\eta}(t)\}\mu_i+$$

$$+\sum_{i,j}E\{\boldsymbol{\eta}(t)'\left[\Psi_{i,j}(\theta_j(t))+\Psi_{i,j}(\theta_j(t))'+\Psi_{i,j}(\theta_j(t))'\Psi_{i,j}(\theta_j(t))\right]\boldsymbol{\eta}(t)\}\mu_j.$$

Using the independence of the random process $\theta_i(t)$ and Assumption 6.1, we can further write

$$\frac{d}{dt}E\{\|\boldsymbol{\eta}(t)\|^2\}=\mu\sum_{i,j}E\{\boldsymbol{\eta}(t)'\boldsymbol{H}\boldsymbol{\eta}(t)\},$$

where

$$\boldsymbol{H}=E\{\Phi_{i,j}(\theta_i(t))+\Phi_{i,j}(\theta_i(t))'+\Phi_{i,j}(\theta_i(t))'\Phi_{i,j}(\theta_i(t))+$$
$$+\Psi_{i,j}(\theta_j(t))+\Psi_{i,j}(\theta_j(t))'+\Psi_{i,j}(\theta_j(t))'\Psi_{i,j}(\theta_j(t))\}.$$

Using Assumption 6.1, we have

$$\Phi_{i,j}(\theta_i(t))+\Phi_{i,j}(\theta_i(t))'=(1-\lambda)\begin{cases}-2 & \text{at entry } (i,j)(i,j)\\ \chi_{\{\theta_i(t)=l\}} & \text{at entries } (i,j)(l,j) \text{ and } (l,j)(i,j)\ l\in\mathcal{N}_i, l\neq j\\ 0 & \text{at all other entries}\end{cases}$$

$$\Phi_{i,j}(\theta_i(t))'\Phi_{i,j}(\theta_i(t))=(1-\lambda)^2\begin{cases}1 & \text{at entry } (i,j)(i,j)\\ -\chi_{\{\theta_i(t)=l\}} & \text{at entries } (i,j)(l,j) \text{ and } (l,j)(i,j)\ l\in\mathcal{N}_i, l\neq j\\ \chi_{\{\theta_i(t)=l\}} & \text{at entries } (l,j)(l,j)\ l\in\mathcal{N}_i, l\neq j\\ 0 & \text{at all other entries}\end{cases}$$

$$\Psi_{i,j}(\theta_j(t))+\Psi_{i,j}(\theta_j(t))'=(1-\lambda)\begin{cases}-2 & \text{at entry } (i,j)(i,j)\\ \chi_{\{\theta_j(t)=l\}} & \text{at entries } (i,j)(i,l) \text{ and } (i,l)(i,j)\ l\in\mathcal{N}_j, l\neq i\\ 0 & \text{at all other entries}\end{cases}$$

$$\Psi_{i,j}(\theta_j(t))'\Psi_{i,j}(\theta_j(t))=(1-\lambda)^2\begin{cases}1 & \text{at entry } (i,j)(i,j)\\ -\chi_{\{\theta_j(t)=l\}} & \text{at entries } (i,j)(i,l) \text{ and } (i,l)(i,j)\ l\in\mathcal{N}_j, l\neq i\\ \chi_{\{\theta_j(t)=l\}} & \text{at entries } (i,l)(i,l)\ l\in\mathcal{N}_j, l\neq i\\ 0 & \text{at all other entries}\end{cases}$$

or

$$E\{\Phi_{i,j}(\theta_i(t))+\Phi_{i,j}(\theta_i(t))'\}=(1-\lambda)\begin{cases}-2 & \text{at entry } (i,j)(i,j)\\ \frac{1}{n-1} & \text{at entries } (i,j)(l,j) \text{ and } (l,j)(i,j)\ l\in\mathcal{N}_i, l\neq j\\ 0 & \text{at all other entries}\end{cases}$$

$$E\{\Phi_{i,j}(\theta_i(t))'\Phi_{i,j}(\theta_i(t))\}=(1-\lambda)\begin{cases}1 & \text{at entry } (i,j)(i,j)\\ -\frac{1}{n-1} & \text{at entries } (i,j)(l,j) \text{ and } (l,j)(i,j)\ l\in\mathcal{N}_i, l\neq j\\ \frac{1}{n-1} & \text{at entries } (l,j)(l,j)\ l\in\mathcal{N}_i, l\neq j\\ 0 & \text{at all other entries}\end{cases}$$

$$E\{\Psi_{i,j}(\theta_j(t))+\Psi_{i,j}(\theta_j(t))'\}=(1-\lambda)\begin{cases}-2 & \text{at entry } (i,j)(i,j)\\ \frac{1}{n-1} & \text{at entries } (i,j)(i,l) \text{ and } (i,l)(i,j)\ l\in\mathcal{N}_j, l\neq i\\ 0 & \text{at all other entries}\end{cases}$$

$$E\{\Psi_{i,j}(\theta_j(t))'\Psi_{i,j}(\theta_j(t))\} = (1-\lambda)^2 \begin{cases} 1 & \text{at entry } (i,j)(i,j) \\ -\frac{1}{n-1} & \text{at entries } (i,j)(i,l) \text{ and } (i,l)(i,j)\ l \in \mathcal{N}_j, l \neq i \\ \frac{1}{n-1} & \text{at entries } (i,l)(i,l)\ l \in \mathcal{N}_j, l \neq i \\ 0 & \text{at all other entries.} \end{cases}$$

Summing up the above matrices we obtain that $\boldsymbol{H}$ is a symmetric matrix that has as diagonal elements quantities of the form

$$\left[-4(1-\lambda)+(1-\lambda)^2\left(2+\frac{2n-4}{n-1}\right)\right]\mu$$

and the off-diagonal, non-zero entries are given by

$$\lambda(1-\lambda)\frac{2}{n-1}\mu.$$

Counting the off-diagonal entries on a row we obtaine the same result as in the case of the first moment. Namely, the number of non-zero and off-diagonal elements on each row is $2(n-2)$. Also note that the diagonal elements are negative and that the off-diagonal and non-zero elements are positive for any $n \geq 2$. Therefore each row of $\boldsymbol{H}$ sums up to the same value and consequently $\boldsymbol{H}$ has an eigenvalue

$$\alpha_2 = \left[-4(1-\lambda)+(1-\lambda)^2\left(2+\frac{2n-4}{n-1}\right)\right]\mu + 2(n-2)\lambda(1-\lambda)\frac{2}{n-1}\mu = -\frac{2(1-\lambda^2)\mu}{n-1},$$

corresponding to eigenvector $\mathbb{1}_{\bar{n}}$. Note that α_2 is negative for $0 \leq \lambda < 1$ and $n \geq 2$. In addition, by Gershgorin' theorem (Theorem 7.2.1, page 320, [4]), we have that all eigenvalues belong to the circle centered at $\left[-4(1-\lambda)+(1-\lambda)^2\left(2+\frac{2n-4}{n-1}\right)\right]\mu$ with radius $2(n-2)\lambda(1-\lambda)\frac{2}{n-1}\mu$ and therefore the eigenvalue α_2 dominates the rest of the eigenvalues; eigenvalues that are real due to symmetry. Therefore, we have that

$$\boldsymbol{H} \leq \alpha_2 I,$$

and consequently

$$\frac{d}{dt}E\{\|\boldsymbol{\eta}(t)\|^2\} \leq \alpha_2 E\{\|\boldsymbol{\eta}(t)\|^2\}.$$

We can further write that

$$E\{\|\boldsymbol{\eta}(t)\|^2\} \leq e^{\alpha_2 t}E\{\|\boldsymbol{\eta}(t_0)\|^2\},$$

from where the result follows. ■

Remark 6.1: As expected, the eigenvalues α_1 and α_2 approach zero as n approaches infinity, and therefore the rate of converges decreases. Interestingly, in both the first and the second moment analysis, we observe that the minimum values of α_1 and α_2 are attained for $\lambda = 0$, that is when an awaken agent never picks its own value, but the value of a neighbor.

VII. The generalized gossip consensus algorithm for particular convex metric spaces

In this section we present several instances of the gossip algorithm for particular examples of convex metric spaces. We consider three cases for $\mathcal{X}$: the set of real numbers, the set of compact intervals and the set of discrete random variables. We endow each of these sets with a metric d and convex structure ψ in order to form convex metric spaces. We show the particular form the generalized gossip algorithm takes for these convex metric spaces, and give some numerical simulations of these algorithms.

A. The set of real numbers

Let $\mathcal{X} = \mathbb{R}$ and consider as metric the standard Euclidean norm, that is $d(x,y) = \|x-y\|_2$, for any $x,y \in \mathbb{R}$. A natural convex structure on $\mathbb{R}$ is given by

$$\psi(x,y,\lambda) = \lambda x + (1-\lambda)y, \ \ \forall x,y \in \mathbb{R}, \lambda \in [0,1]. \tag{24}$$

Indeed since for a point $z \in \mathbb{R}$

$$d(z,\psi(x,y,\lambda)) = \|z-(\lambda x+(1-\lambda)y)\|_2 = \|\lambda(z-x)+(1-\lambda)(z-y)\|_2 \leq$$
$$\leq \lambda\|z-x\|_2 + (1-\lambda)\|z-y\|_2 = \lambda d(z,x) + (1-\lambda)d(z,y),$$

ψ is a convex structure. Therefore $(\mathbb{R}, \|\cdot\|_2, \psi)$ is a convex metric space. For this particular convex metric space, the generalized randomized consensus algorithm takes the following form

Algorithm 1: Randomized gossip algorithm on $\mathbb{R}$

Input: $x_i(0)$, λ_i, $p_{i,j}$
for *each counting instant t_i of N_i* **do**
Agent i enters update mode and picks a neighbor j with probability $p_{i,j}$;
Agent i updates its state according to

$$x_i(t_i^+) = \lambda_i x_i(t_i) + (1-\lambda_i)x_j(t_i);$$

Agent i enters sleep mode;

Note that this algorithm is exactly the randomized gossip algorithm for solving the consensus problem, which was studied in [2].

B. The set of compact intervals

Let $\mathcal{X}$ be the family of closed intervals, that is $\mathcal{X} = \{[a,b] \mid -\infty < a \leq b < \infty\}$. For $x_i = [a_i, b_i]$, $x_j = [a_j, b_j]$ and $\lambda \in [0,1]$, we define a mapping ψ by $\psi(x_i, x_j, \lambda) = [\lambda a_i + (1-\lambda)a_j, \lambda b_i + (1-\lambda)b_j]$ and use as metric the Hausdorff distance given by

$$d(x_i, x_j) = \max\{|a_i - a_j|, |b_i - b_j|\}.$$

Then, as shown in [24], $(\mathcal{X}, d, \psi)$ is a convex metric space. For this convex metric space, the randomized gossip consensus algorithm becomes,

Algorithm 2: Randomized gossip algorithm on a set of compact intervals

Input: $x_i(0)$, λ_i, $p_{i,j}$
for *each counting instant t_i of N_i* **do**
Agent i enters update mode and picks a neighbor j with probability $p_{i,j}$;
Agent i updates its state according to

$$x_i(t_i^+) = [\lambda_i a_i(t_i) + (1-\lambda_i)a_j(t_i), \lambda_i b_i(t_i) + (1-\lambda_i)b_j(t_i)];$$

Agent i enters sleep mode;

C. The set of discrete random variables

In this section we apply our algorithm on a particular convex metric space that allows us the obtain a probabilistic algorithm for reaching consensus on discrete sets.

Let $S = \{s_1, s_2, \dots, s_m\}$ be a finite and countable set of real numbers and let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. We denote by $\mathcal{X}$ the space of discrete measurable functions (random variable) on $(\Omega, \mathcal{F}, \mathcal{P})$ with values in S.

We introduce the operator $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, defined as

$$d(X,Y) = E_{\mathcal{P}}[\rho(X,Y)], \tag{25}$$

where $\rho : \mathbb{R} \times \mathbb{R} \to \{0,1\}$ is the discrete metric, i.e.

$$\rho(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

and the expectation is taken with respect to the measure $\mathcal{P}$. It is not difficult to note that the operator d can also be written as $d(X,Y) = E[\mathbb{1}_{\{X \neq Y\}}] = Pr(X \neq Y)$, where $\mathbb{1}_{\{X \neq Y\}}$ is the indicator function of the event $\{X \neq Y\}$.

We note that for all $X, Y, Z \in \mathcal{X}$, the operator d satisfies the following properties

(a) $d(X,Y) = 0$ if and only if $X = Y$ with probability one,
(b) $d(X,Z) + d(Y,Z) \geq d(X,Y)$ with probability one,
(c) $d(X,Y) = d(Y,X)$,
(d) $d(X,Y) \geq 0$,

and therefore is a metric on $\mathcal{X}$. The set $\mathcal{X}$ together with the operator d define the *metric space* $(\mathcal{X}, d)$.

Let $\gamma \in \{1,2\}$ be an independent random variable defined on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with probability mass function $Pr(\gamma = 1) = \lambda$ and $Pr(\gamma = 2) = 1 - \lambda$, where $\lambda \in [0,1]$. We define the mapping $\psi : \mathcal{X} \times \mathcal{X} \times [0,1] \to \mathcal{X}$ given by

$$\psi(X_1, X_2, \lambda) = \mathbb{1}_{\{\gamma=1\}} X_1 + \mathbb{1}_{\{\gamma=2\}} X_2, \ \forall X_1, X_2 \in \mathcal{X}, \lambda \in [0,1]. \tag{26}$$

Proposition 7.1: The mapping ψ is a convex structure on $\mathcal{X}$.

Proof: For any $U, X_1, X_2 \in \mathcal{X}$ and $\lambda \in [0,1]$ we have

$$\begin{aligned} d(U, \psi(X_1, X_2, \lambda)) &= E[\rho(U, \psi(X_1, X_2, \lambda))] = E[E[\rho(U, \psi(X_1, X_2, \lambda)) | U, X_1, X_2]] = \\ &= E[E[\rho(U, \mathbb{1}_{\{\gamma=1\}} X_1 + \mathbb{1}_{\{\gamma=2\}} X_2)] | U, X_1, X_2] = E[\lambda \rho(U, X_1) + (1-\lambda)\rho(U, X_2)] = \\ &= \lambda d(U, X_1) + (1-\lambda) d(U, X_2). \end{aligned}$$

■

From the above proposition it follows that $(\mathcal{X}, d, \psi)$ is a *convex metric space*. For this particular convex metric space the randomized consensus algorithm is summarized in what follows.

Let us now take a closer look to the probabilistic model of the above algorithm. Let $\theta_i(t)$ be independent random variables with probability distribution $Pr(\theta_i(t) = j) = p_{i,j}$ for $j \in \mathcal{N}_i$ and for all t, with $\sum_{j \in \mathcal{N}_i} p_{i,j} = 1$. In addition, let $\gamma_i(t)$ be a set of independent random variable such that $Pr(\gamma_i(t) = 1) = \lambda_i$ and $Pr(\gamma_i(t) = 2) = 1 - \lambda_i$, for all t. Then according to the generalized gossip algorithm, at each time instant t_i at which the counter $N_i(t)$ updates its value, agent i updates its value according to the formula

$$x_i(t_i^+) = \chi_{\{\gamma_i(t_i)=1\}} x_i(t_i) + \chi_{\{\gamma_i(t_i)=2\}} \sum_{j \in \mathcal{N}_i} \chi_{\{\theta_i(t_i)=j\}} x_j(t_i). \tag{27}$$

Let $(\bar{\Omega}, \bar{\mathcal{P}}, \bar{\mathcal{F}})$ be a probability space, with $\bar{\mathcal{F}}_t$ a filtration of $\bar{\mathcal{F}}$ given by

$$\bar{\mathcal{F}}_t = \sigma(x_i(s), N_i(s), \gamma_i(s), \theta_i(s), 0 \leq s < t, i = 1, \dots, n),$$

Algorithm 3: Randomized gossip algorithm on countable, finite sets

Input: $x_i(0)$, λ_i, $p_{i,j}$
for *each counting instant* t_i *of* N_i **do**
Agent i enters update mode and picks a neighbor j with probability $p_{i,j}$;
Agent i updates its state according to

$$x_i(t_i^+) = \begin{cases} x_i(t_i) & \text{with probability } \lambda_i \\ x_j(t_i) & \text{with probability } 1-\lambda_i \end{cases}$$

Agent i enters sleep mode;

where we used σ as a symbol for sigma algebra. By (27), it follows $x_i(t)$ is adapted to the filtration $\bar{\mathcal{F}}_t$. Let us now consider the filtration $\mathcal{F}_t^N = \sigma(N_i(s), 0 \leq s < t, i = 1, \ldots, n)$, induced by the Poisson counters $N_i(t)$. In order to accommodate the contribution of the Poisson counters to the probability model of the algorithm, we must refine the metric proposed in (25). As a consequence, at each time instant, the distance between agents is given by

$$d(x_i(t), x_j(t)) = E_{\bar{\mathcal{P}}}[\rho(x_i(t), x_j(t))|\mathcal{F}_t^N],$$

where the expectation is taken with respect to the measure $\bar{\mathcal{P}}$. Consequently, $d(x_i(t), x_j(t))$ is measurable with respect to the sigma-algebra $\mathcal{F}_t^N$.

D. Numerical simulations

In this subsection we present numerical simulations of the generalized gossip algorithm in the case of the three convex metric spaces previously mentioned. We consider a network of $n = 20$ nodes, each of which is equipped with a Poisson counter with rate μ derived uniformly and independently from the interval $(0, 6]$. The Poisson counters are independent among agents. Furthermore each agent will be equipped with a convex parameter λ which is chosen uniformly and independently from $[0, 1)$. At every ignition time each agent will connect with any other agent with probability $p = \frac{1}{n-1}$. For each the three convex metric spaces we present two figures: the first figures show the values of the states, while the second figure depicts upper bound on of the distances between the agents' states, that is the quantities $\eta_{i,j}(t)$. Our focus is on showing that the vector of distances converge to zero and that the states converge to the same value and therefore to simplify the figures' depiction, all quantities are represented using the black color.

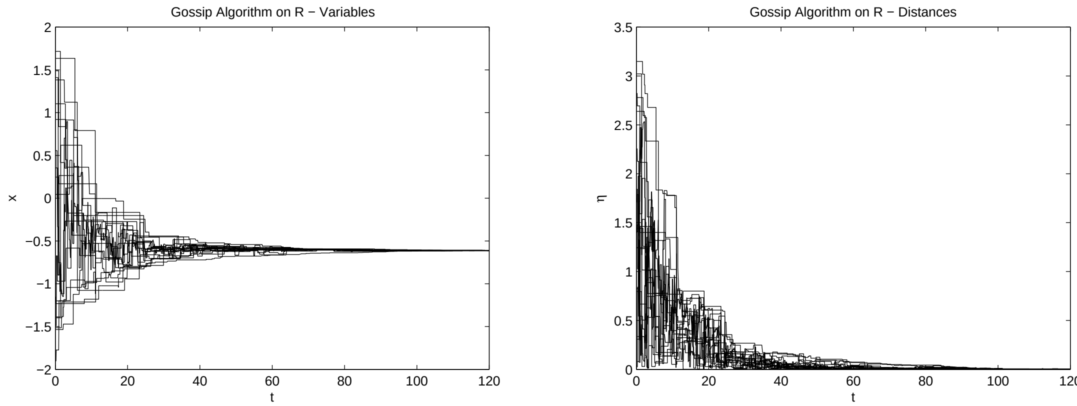
Figure 1 shows a realization of the generalized consensus algorithm in the case of the convex metric space defined on the set of real numbers. We assume that the agents initialize their values uniformly from the interval $[-2, 2]$. As expected the distances between the states of the agents converge to zero and the states converge to the same value.

Figure 2 shows a realization of the generalized gossip algorithm in the case the convex metric space is defined on the set of closed intervals. We present a sample path of the actual values of the states (which are intervals of the form $[a_i, b_i]$) together with the corresponding bounds on the distances between the states. We consider the initial intervals to be initialized in $[-2, 2]$.

Figure 3 shows a realization of the generalized gossip algorithm in the case the set $\mathcal{X}$ is given by the set of discrete random variables. As in the case of the set of real numbers, the agents initialize their values uniformly from the interval $[-2, 2]$, however their values will belong to the set $\{x_1(0), x_2(0), \ldots, x_n(0)\}$ for all time instants. As a consequences, both the states and the distances oscillate more, but nonetheless the distances converge to zero and the states converge to a common value.

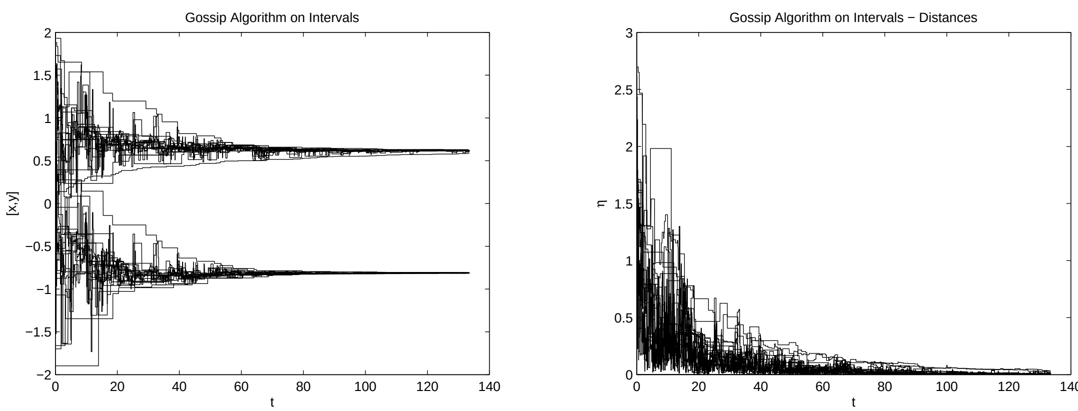
VIII. Conclusions

In this paper we analyzed the convergence properties of a generalized randomized gossip algorithm acting on convex metric spaces. We gave convergence results in almost sure and r^{th} mean sense for the



(a) (b)

Fig. 1: Randomized Gossip Algorithm on $\mathbb{R}$: (a) the values of the states; (b) (upper bounds on the) distances between the states of the agents.



(a) (b)

Fig. 2: Randomized Gossip Algorithm on Closed Intervals: (a) the values of the states (the vertical axis depicts both ends of the intervals $[a_i(t), b_i(t)]$ which stand for the agents' states); (b) (upper bounds on the) distances between the states of the agents.

distances between the states of the agents. Under specific assumptions on the communication topology, we computed explicitly estimates of the rate of convergence for the first and second moments of the distances between the agents. Additionally, we introduces instances of the generalized gossip algorithm for three particular convex metric spaces and presented numerical simulations of the algorithm.

References

[1] V. Borkar and P. Varaya. Asymptotic agreement in distributed estimation. *IEEE Trans. Autom. Control*, AC-27(3):650–655, Jun 1982.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14(SI):2508–2530, 2006.

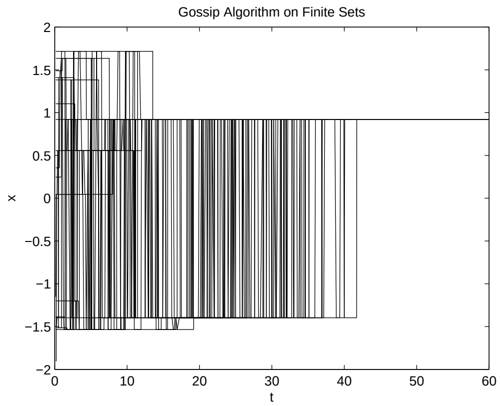
[3] GN. Dunford and J.T. Schwartz. *Linear operators, Part 1.* Interscience, New York, 1958.

[4] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, USA, third edition edition, 1996.

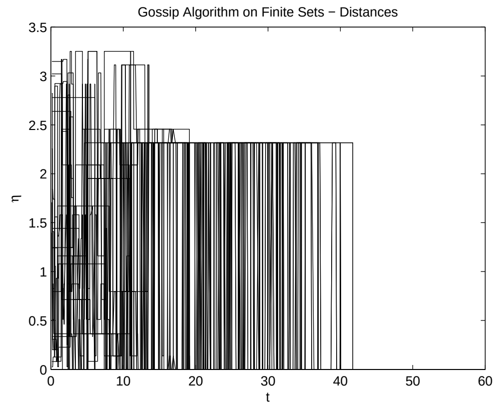
[5] G. Grimmet and D. Stirzaker. *Probability and Random Processes*. Oxford University Press, third edition edition, 2001.

[6] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Trans. Autom. Control*, 50(11):1867–1872, 2005.

[7] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor. *IEEE Trans. Autom. Control*, 48(6):998–1001, Jun 2004.



(a)



(b)

Fig. 3: Randomized Gossip Algorithm on Discrete Finite Sets: (a) the values of the states; (b) (upper bounds on the) distances between the states of the agents.

[8] I. Matei and J. Baras. The asymptotic consensus problem on convex metric spaces. *Proceedings of the 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems(NecSys 2010)*, pages 287–292, September 2010.

[9] I. Matei and J. Baras. The geometry of a probabilistic consensus of opinion algorithm. *Proceedings of the 2011 American Control Conference*, pages 2198–2203, June 2011.

[10] I. Matei, N. Martins, and J. Baras. Almost sure convergence to consensus in markovian random graphs. *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 3535–3540, Dec 2008.

[11] I. Matei, N. Martins, and J. Baras. Consensus problems with directed markovian communication patterns. *Proceedings of the 2009 IEEE American Control Conference*, pages 1298–1303, Jun 2009.

[12] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics, 2000.

[13] L. Moreau. Stability of multi-agents systems with time-dependent communication links. *IEEE Trans. Autom. Control*, 50(2):169–182, Feb 2005.

[14] A. Nedic and A. Ozdaglar. Convergence rate for consensus with delays. *Journal of Global Optimization*, 47(3):437–456, Jul 2010.

[15] A. Nedic, A. Ozdaglar, and P.A. Parrilo. Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Autom. Control*, 55(4):922–938, Apr 2010.

[16] R. Olfati-Saber and R.M. Murray. Consensus protocols for networks of dynamic agents. *Proceedings of the 2003 IEEE American Control Conference*, pages 951–956, 2003.

[17] R. Olfati-Saber and R.M. Murray. Consensus problem in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, 49(9):1520–1533, Sept 2004.

[18] M. Porfiri and D.J. Stilwell. Consensus seeking over random directed weighted graphs. *IEEE Trans. Autom. Control*, 52(9):1767–1773, Sept 2007.

[19] W. Ren and R.W. Beard. Consensus seeking in multi-agents systems under dynamically changing interaction topologies. *IEEE Trans. Autom. Control*, 50(5):655–661, May 2005.

[20] A. Tahbaz Salehi and A. Jadbabaie. Necessary and sufficient conditions for consensus over random networks. *IEEE Trans. Autom. Control*, 53(3):791–795, Apr 2008.

[21] A. Tahbaz Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Trans. Autom. Control*, 55(1):225–230, Jan 2010.

[22] D. Shah. Gossip algorithms. *Foundations and Trends in Networking*, 3(1):1–125, 2008.

[23] B.K. Sharma and C.L. Dewangan. Fixed point theorem in convex metric space. *Novi Sad Journal of Mathematics*, 25(1):9–18, 1995.

[24] W. Takahashi. A convexity in metric mpace and non-expansive mappings i. *Kodai Math. Sem. Rep.*, 22:142–149, 1970.

[25] J.N. Tsitsiklis. Problems in decentralized decision making and computation. *Ph.D. dissertation, Dept. Electr. Eng.*, Nov 1984.

[26] J.N. Tsitsiklis, D.P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Trans. Autom. Control*, 31(9):803–812, Sept 1986.